Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **451,** 3964–3974 (2015) doi:10.1093/mnras/stv1228

Testing spherical evolution for modelling void abundances

Ixandra Achitouv,[1,2]★ Mark Neyrinck[3] and Aseem Paranjape[4]

[1]*Ludwig-Maximilians-Universitat München, Universitäts-Sternwarte München, Scheinerstr. 1, D-81679 München, Germany*
[2]*Excellence Cluster Universe, Boltzmannstr. 2, D-85748 Garching bei München, Germany*
[3]*Department of Physics and Astronomy, The Johns Hopkins University, Baltimore, MD 21218, USA*
[4]*Department of Physics, Institute for Astronomy, ETH Zürich, Wolfgang-Pauli-Strasse 27, CH-8093 Zürich, Switzerland*

Accepted 2015 May 28. Received 2015 May 19; in original form 2013 September 23

ABSTRACT

We compare analytical predictions of void volume functions to those measured from *N*-body simulations, detecting voids with the ZOBOV void finder. We push to very small, non-linear voids, below few Mpc radius, by considering the unsampled dark matter density field. We also study the case where voids are identified using haloes. We develop analytical formula for the void abundance of both the excursion set approach and the peaks formalism. These formulas are valid for random walks smoothed with a top-hat filter in real space, with a large class of realistic barrier models. We test the extent to which the spherical evolution approximation, which forms the basis of the analytical predictions, models the highly aspherical voids that occur in the cosmic web, and are found by a watershed-based algorithm such as ZOBOV. We show that the volume function returned by ZOBOV is quite sensitive to the choice of treatment of subvoids, a fact that has not been appreciated previously. For reasonable choices of subvoid exclusion, we find that the Lagrangian density $\delta_{\rm v}$ of the ZOBOV voids – which is predicted to be a constant $\delta_{\rm v} \approx -2.7$ in the spherical evolution model – is different from the predicted value, showing substantial scatter and scale dependence. This result applies to voids identified at $z = 0$ with effective radius between 1 and $10\,h^{-1}$ Mpc. Our analytical approximations are flexible enough to give a good description of the resulting volume function; however, this happens for choices of parameter values that are different from those suggested by the spherical evolution assumption. We conclude that analytical models for voids must move away from the spherical approximation in order to be applied successfully to observations, and we discuss some possible ways forward.

Key words: methods: analytical – methods: numerical – cosmology: theory – large-scale structure of Universe.

1 INTRODUCTION

A visually striking aspect of all galaxy surveys to date is the presence of large, nearly empty regions known as *voids* (Kirshner et al. 1981; Kauffmann & Fairall 1991; Hoyle & Vogeley 2002, 2004; Croton et al. 2004; Patiri et al. 2006b; Pan et al. 2012; Sutter et al. 2012). There has been considerable interest in characterizing the observable properties of voids and understanding their origin and dynamics (Hoffman, Salpeter & Wasserman 1983; Dubinski et al. 1993; van de Weygaert & van Kampen 1993; Sahni, Sathyaprakah & Shandarin 1994; Colberg et al. 2005; Patiri, Betancort-Rijo & Prada 2006a; van de Weygaert & Platen 2011). While a void can be defined in many ways (see e.g. Colberg et al. 2008, and references therein), the basic picture of a large, underdense, expanding region (Fillmore & Goldreich 1984; Bertschinger 1985) has stood the test of time. Typical void sizes depend on the type of galaxy used to define them; e.g. in the main sample of the Sloan Digital Sky Survey, they can range from $\sim$15 to $\sim 30\,h^{-1}$ Mpc (e.g. Pan et al. 2012; Sutter et al. 2012), while there are also examples of voids as large as $\sim 100\,h^{-1}$ Mpc (e.g. Granett, Neyrinck & Szapudi 2008).

The presence of voids in galaxy surveys leads to many questions: whether galaxies that reside in void environments are special (Goldberg et al. 2005; Hoyle et al. 2005); whether large, deep voids are a challenge to structure formation in Λ cold dark matter (ΛCDM) cosmologies (Blumenthal et al. 1992; Hunt & Sarkar 2010), or whether they are a natural consequence of the well-understood dynamics of CDM (Tinker & Conroy 2009); whether voids can then be used as a cosmological tool to distinguish between models (Ryden 1995; Park & Lee 2007; Kamionkowski, Verde & Jimenez 2009; Lam, Sheth & Desjacques 2009; Biswas, Alizadeh & Wandelt 2010; Lavaux & Wandelt 2010, 2012; D'Amico et al. 2011; Hamaus et al.

★ E-mail: achitouv@usm.lmu.de

© 2015 The Authors
Published by Oxford University Press on behalf of the Royal Astronomical Society

2014; Melchior et al. 2014; Pisani et al. 2014); and whether their dynamics and statistics can be modelled analytically (Sheth & van de Weygaert 2004; Furlanetto & Piran 2006; D'Aloisio & Furlanetto 2007).

Analytical models for isolated voids have been well studied in the literature for decades (Bertschinger 1985; Blumenthal et al. 1992). A major advance in their *statistical* modelling was presented by Sheth & van de Weygaert (2004, hereafter SvdW), who demonstrated that voids obey a hierarchy similar to that of haloes. In particular, their analysis led to a prediction for the size distribution of voids based on the excursion set approach (Press & Schechter 1974; Epstein 1983; Bond et al. 1991). Essentially, voids are modelled as regions that are initially underdense enough to reach shell-crossing by the present epoch. The SvdW analysis had three shortcomings, however: (a) it was based on an excursion set model using random walks in the smoothed density field with *uncorrelated* rather than correlated steps; (b) it was entirely based upon the *initial* or 'Lagrangian' dark matter (DM) density; and (c) the intrinsic averaging of the excursion set walks (on randomly selected position) was not taken into account. Recently, these shortcomings were overcome. In Paranjape, Lam & Sheth (2012) the SvdW treatment was modified to account for both correlated steps in the random walks [which arise when using smoothing filters such as the real-space top-hat (SX filter)] as well as the fact that voids are identified in the evolved 'Eulerian' field. In Achitouv et al. (2013, 2014), it was shown that the consistency of the excursion set framework is preserved once the barrier threshold is extended to stochastic modelling, and also shortcoming (a) was solved using an alternative path integral approach that we apply to voids in this work.

Despite these improvements, excursion set void models cannot be directly compared with the distribution of observed galaxy voids. This is because these models are meant to describe voids in the DM, whereas the galaxies used to define voids observationally are biased tracers of DM. Furlanetto & Piran (2006) showed how galaxies can be included in the analysis by combining the SvdW excursion set calculation with the halo model (Peacock & Smith 2000; Seljak 2000). As expected when using biased tracers, this increases the sizes of voids in a manner that is correlated with galaxy type (e.g. more luminous galaxies define larger voids on average). The size distributions of observed galaxy voids, e.g. those presented by Pan et al. (2012) or Sutter et al. (2012), should therefore be compared with predictions such as those of Furlanetto & Piran (2006) and not with SvdW.

Before doing this, however, it is important to ask whether the SvdW model (or the improved version suggested by Paranjape et al. 2012) gives a good description of voids identified in the DM density itself, which is possible in N-body simulations (Colberg et al. 2005; Jennings, Li & Hu 2013), and is one of the primary goals of this paper. Jennings et al. (2013) recently compared a modified version of the SvdW predictions (where the volume conservation is enforced) to the results of a void finder specifically built to identify spherical underdensities. Although it is plausible that this is the correct way of comparing the SvdW predictions with measurements, it ignores the highly aspherical, polyhedral shape that initial underdensities develop into as they form voids, and the unrealistic void volumes defined by the sharp-k (SK) filter used in SvdW. It is therefore interesting to ask whether the voids identified by popular algorithms (we use ZOBOV below) can be incorporated in an appropriate analytical framework that goes beyond the approximation of spherical evolution. Moreover, from a physical point of view, one also expects that voids tend to form near minima of the initial density field (see Colberg et al. 2005, who demonstrated this in N-body simulations), and it is then interesting to ask whether including a peaks constraint (Bardeen et al. 1986) in the excursion set calculation improves the comparison.

The plan of the paper is as follows. In Section 2 we present analytical results for the void volume function based on path-integral calculations within the excursion set approach which we test against Monte Carlo simulations of random walks. We also present the results of including the peaks constraint in such a calculation, and extend both these results to the case of stochastic and scale-dependent void-formation thresholds. In Section 3 we turn to voids identified in N-body simulations. We describe the simulations and discuss the ZOBOV void finder. In particular, we explore the sensitivity of the latter to the choice of treatment of substructures within the identified voids, and compare our analytical results with the ZOBOV voids. We also study the effect of sampling the DM particles by considering voids identified with haloes. The effect of the biasing is also tested by comparing this result with randomly selected DM particles. In Section 4 we check whether ZOBOV voids are consistent with the assumptions of the spherical evolution model by measuring the initial overdensity at an appropriately defined void centre. In Section 5, to test the sensitivity of our results to the particular void finder, we repeat some of our comparisons for voids found using a spherical underdensity finder (e.g. Jennings et al. 2013). We conclude in Section 6 with a summary of our results and prospects for future work.

2 THE VOID HIERARCHY

If the statistics of voids carry cosmological information, then a successful theory should be able to predict void properties directly from the initial conditions once the cosmological background is known. The most naive idea is to link the site of a void to an underdense region in Lagrangian space. Assuming this initial depression evolves decoupled from the surrounding shear field, and is approximately spherical, then SvdW have shown that the linear critical underdensity required to form a void at $z = 0$ is approximately $\delta_{\rm v} = -2.7$ in an Einstein–de Sitter universe. Unlike haloes, voids expand over time and repel matter. A spherical evolution model predicts that the Eulerian radius ($R_{\rm E}$) of a void is $R_{\rm E} \sim 1.7R$, with R its Lagrangian size. This deterministic mapping is more linear compared to the collapse of protohaloes (which contract by a factor of $\sim$5.8). The density within the void is $\Delta_{\rm v}(z = 0) \sim -0.8$.

This rather simple analytical model is the building block which allows to pass from the statistical properties of voids in the matter density field to the statistical properties using biased tracers such as galaxies (Furlanetto & Piran 2006). The linear spherical threshold $\delta_{\rm v}$ can be used to predict the site of void formation from the Lagrangian field. However, the dynamics of voids are subject to an additional, void-in-cloud effect (SvdW). This occurs in a region which is collapsing (or has collapsed) on a large scale R_1, but is underdense on a smaller scale R_2. In what follows, we make predictions for the void abundance using a realistic volume prediction within the standard excursion set approach, and using a modified peak-excursion set approach. For both cases we also extend the spherical threshold to more general class of barriers defined by a Gaussian and a log-normal distribution.

2.1 Excursion set approach

The standard excursion set theory (Bond et al. 1991) is a useful framework to compute the abundance of DM haloes, and can also be

applied to voids. The key assumption is to equate the volume fraction in voids of radius R to an appropriate first-crossing distribution:

$$V\frac{\mathrm{d}n}{\mathrm{d}\ln R} = f(\sigma)\left|\frac{\mathrm{d}\ln\sigma}{\mathrm{d}\ln R}\right|, \quad (1)$$

where $f(\sigma) \equiv 2\sigma^2\mathcal{F}(\sigma)$ is the so-called multiplicity function and $\mathcal{F} \equiv \mathrm{d}F/\mathrm{d}S$ is the derivative of the volume fraction arising from the first-crossing problem. Let us denote the probability density that an overdensity smoothed on a scale $R(S)$ is below a critical threshold B by $\Pi(\delta, S(R))$. We have

$$F(S(R)) = -\int_{-\infty}^{B}\Pi(\delta, S(R))\,\mathrm{d}\delta + C, \quad (2)$$

where C is a constant independent of the scale R, and S is the variance of the associated field:

$$S \equiv \langle\delta^2(R)\rangle \equiv \sigma^2 = \frac{1}{2\pi^2}\int \mathrm{d}k\,k^2P(k)\tilde{W}^2(k, R). \quad (3)$$

Once the filter $W(k, R)$ is specified, there is a one-to-one mapping between the scale R and the variance S.

This formalism can be extended in the void hierarchy where voids are characterized by their volume rather than the mass they encapsulate.

Another common quality of the excursion set theory is that the so-called cloud-in-cloud issue is solved: a collapsing structure cannot be embedded in a larger one which would lead to miscounting the number of haloes. This issue appears if the smoothed overdensity crosses the threshold at multiple smoothing scales and can be treated by adding an absorbing boundary condition: $\Pi(\delta = B, S) = 0$ such that the largest scale defines the mass $M(R)$. In the case of voids, the void-in-cloud process, describing collapsing voids, is important to take into account, as described in SvdW for a SK filter. Therefore we must distinguish between the barrier associated with haloes (denoted B_h) and the one associated with voids (denoted B_v). Thus all the game is to compute the $\Pi(\delta, S)$ under the condition that $\Pi(\delta, S = 0) = \delta_\mathrm{D}(\delta)$, $\Pi(\delta = B_\mathrm{h}, S) = \Pi(\delta = B_\mathrm{v}) = 0$ and compute the first-crossing $F(S|\delta(S') < B_\mathrm{h}(S'))$ with $S' < S$. For a SK filter and a constant barrier (e.g. $B_\mathrm{v} = -2.7$, $B_\mathrm{h} = 1.686$) the solution of this system is given in SvdW. The extension to a linear moving barrier of the same slope ($B_\mathrm{v} = \delta_\mathrm{v} - \beta S$, $B_\mathrm{h} = \delta_\mathrm{c} - \beta S$) can be found in appendix C of SvdW, while the extension to positive slope has been worked out in appendix A of Furlanetto & Piran (2006). Note that for the halo barrier, ellipsoidal collapse predicts a positive slope. However, we will see in Section 4 that for the void threshold, it seems that a negative slope is in better agreement with the Lagrangian barrier. However, before jumping to the barrier criteria we should emphasize that all those predictions hold for a particular type of filter, a top-hat in Fourier space (SK). The volume encapsulated by such a filter is given by[1]

$$V_\mathrm{SK}(R) = \int \mathrm{d}^3R\;W_\mathrm{SK}(R) = 6\pi^2R^3 - 12\pi R^3\int_0^\infty \cos x\,\mathrm{d}x. \quad (4)$$

One could argue that the divergent integral part can be set to zero (see e.g. Lacey & Cole 1993). However, it is more difficult to picture the shape associated with such a void and in addition, this volume is never used in observations or in N-body simulations to define structures. Therefore, if we assume a spherical shell evolution of the void, for consistency, the appropriate filter should be a SX filter, which defines a spherical volume. However, in this case there is no exact analytical solution to the first crossing. One could run Monte Carlo walks (Bond et al. 1991) and solve the exact associated first-crossing, which would be straightforward (Paranjape et al. 2012). Nevertheless, a path-integral approach (Maggiore & Riotto 2010a; Corasaniti & Achitouv 2011b) can be used to compute analytically the correction induced by such a filter (see Musso & Sheth (2012) for alternative methods). The SX filter introduces small corrections to the SK case, which can be computed perturbatively and applied to halo formation in the excursion set framework. The amplitude of this correction is weakly dependent on the smoothing scale, and is set by the linear matter power spectrum. This method has been shown to be very accurate and to converge well: the exact Monte Carlo solution matches the analytical approximation with high accuracy (Corasaniti & Achitouv 2011a; Achitouv et al. 2013). In what follows, we investigate the pertinence of the void-in-cloud effect for realistic halo thresholds in the context of the excursion set theory, and show that a one-barrier threshold is a very good approximation to the exact Monte Carlo solution, providing a simple analytical formula for the SK filter. Finally we extend this prediction to the SX filter. Our results are consistent with previous work by SvdW for SK filtering and Paranjape et al. (2012) for SX filtering. See also Zhang & Hui (2006) and Lam & Sheth (2009) for a complementary approach.

Achitouv et al. (2013) found that within the excursion set framework, any consistent barrier should have an intrinsic scatter due to the randomness of the position that the excursion set theory assumes in order to compute the fraction of collapsed regions. Note that deviations from spherical collapse also contribute to this scatter.[2] Over the range they investigate, they found that a Gaussian barrier with a mean value of $\langle B\rangle(S) = \delta_\mathrm{c} + \beta S$ and rms $\sqrt{D_\mathrm{B}S}$ is consistent with the initial Lagrangian critical overdensity leading to halo formation, and predicts a mass function which is in very good agreement with N-body simulations (e.g. Corasaniti & Achitouv 2011a; Achitouv & Corasaniti 2012a). Therefore, in order to test the void-in-cloud effect on the abundance of void, we assume a realistic barrier for halo and void formation (a diffusive drifting barrier). Note that in this model the random walk performed by the barrier is not correlated with the one performed in δ. See Achitouv et al. (2013) for a discussion on this assumption.

Following Bond et al. (1991), we perform Monte Carlo random walks to solve the first-crossing associated with a generic filter and barrier, and we implement the condition that walks which cross the void barrier on a scale S_1 never cross the halo barrier on a smaller scale $S_2 < S_1$. For the halo barrier, we take $\beta = 0.1, D_\mathrm{B} = 0.4$. Similarly, for the void barrier we consider a barrier with a Gaussian distribution characterized by a mean $\langle B_\mathrm{v}\rangle = \delta_\mathrm{v} - \beta S$ and rms $\sqrt{D_\mathrm{B}S}$, with the same choices $\beta = 0.1, D_\mathrm{B} = 0.4$. Note that this choice of parameters is rather arbitrary, because there is no theoretical prediction for them. We flip the sign of the slope for reasons that will become clear in Section 4.

The Monte Carlo results for the SK filter are shown in Fig. 1. On the left panel, the light-blue histogram shows the Monte Carlo result associated with the two-barrier condition, while the blue dots neglect the void-in-cloud effect. As we can see, the void-in-cloud effect operates at low radius. This effect also depends on the halo threshold as it was discussed in SvdW. For a drifting diffusive barrier and Markovian walks (SK filter), the void-in-cloud effect appears

[1] See also discussion after equation (36) in Maggiore & Riotto (2010a).

[2] The reconstruction of the barrier for the centre of mass (i.e. on the peak of the protohalo) also shows a scatter (Robertson et al. 2009; Achitouv et al. 2013, 2014).

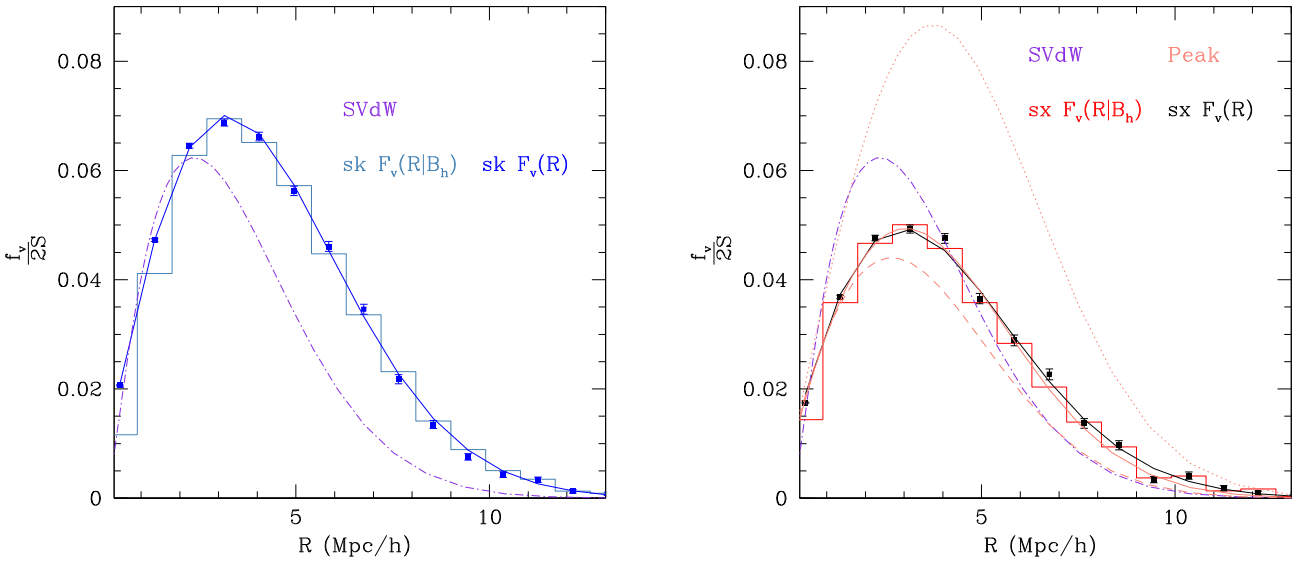

Figure 1. Void multiplicity function as a function of void radius for Markovian random walks (SK) (left-hand panel) and correlated walks (SX) (right-hand panel). The histogram shows the exact solution including the treatment of the void-in-cloud effect while the dots neglect this issue. Blue (left panel) and black (right panel) solid lines are the theoretical predictions of equations (5) and (9), orange lines (right panel), correspond to peak prediction (see text), and the purple dot–dashed line shows the SvdW reference (SK).

at $R < 3\,\mathrm{Mpc}\,h^{-1}$. The blue solid line is the analytical prediction of the excursion set theory for a diffusive drifting void threshold with $\langle B_\mathrm{v}(S)\rangle = \delta_\mathrm{v} - \beta S$ and $\langle B_\mathrm{v}(S_1)B_\mathrm{v}(S_2)\rangle = D_\mathrm{B}\min(S_1, S_2)$. This solution neglects the void-in-cloud effect and is exact for the SK filter. It was computed as in Corasaniti & Achitouv (2011a,b), leading to a Markovian multiplicity function (SK) for voids:

$$f_0(\sigma) = \frac{\delta_\mathrm{v}}{\sigma}\sqrt{\frac{2a}{\pi}}\,\mathrm{e}^{-\frac{a}{2\sigma^2}(\delta_\mathrm{v}-\beta\sigma^2)^2}, \tag{5}$$

with $a = 1/(1 + D_\mathrm{B})$.

The original prediction of SvdW is shown in purple. Note also that equation (5) reproduces with high accuracy SvdW by simply setting $\beta = 0, D_\mathrm{B} = 0$ ($R > 2\,\mathrm{Mpc}\,h^{-1}$). However, in order to have a coherent volume definition, we should consider walks smoothed with a (SX) filter when computing the multiplicity function. In this case, we use the same path integral technique as Maggiore & Riotto (2010a,b), Corasaniti & Achitouv (2011a), and Achitouv & Corasaniti (2012b). Taking into account void-in-void and neglecting void-in-cloud effects, we find that the non-Markovian corrections for a diffusive drifting barrier are

$$f^{m-m}_{1,\beta=0}(\sigma) = \tilde{\kappa}\,\frac{\delta_\mathrm{v}}{\sigma}\sqrt{\frac{2a}{\pi}}\left[\mathrm{e}^{-\frac{a\delta_\mathrm{v}^2}{2\sigma^2}} - \frac{1}{2}\Gamma\left(0, \frac{a\delta_\mathrm{v}^2}{2\sigma^2}\right)\right], \tag{6}$$

$$f^{m-m}_{1,\beta^{(1)}}(\sigma) = a\,\delta_\mathrm{v}\,\beta\left[\tilde{\kappa}\,\mathrm{Erfc}\left(\delta_\mathrm{v}\sqrt{\frac{a}{2\sigma^2}}\right) + f^{m-m}_{1,\beta=0}(\sigma)\right], \tag{7}$$

$$f^{m-m}_{1,\beta^{(2)}}(\sigma) = -a\,\beta\left[\frac{\beta}{2}\sigma^2 f^{m-m}_{1,\beta=0}(\sigma) - \delta_\mathrm{v}\,f^{m-m}_{1,\beta^{(1)}}(\sigma)\right], \tag{8}$$

where $\tilde{\kappa} = a\,\kappa$, and κ is set by the linear matter power spectrum. For a vanilla ΛCDM universe, $\kappa \sim 0.465$, giving the following void total multiplicity function for a SK filter:

$$f_\mathrm{v}(\sigma) = f_0(\sigma) + f^{m-m}_{1,\beta=0}(\sigma) + f^{m-m}_{1,\beta^{(1)}}(\sigma) + f^{m-m}_{1,\beta^{(2)}}(\sigma). \tag{9}$$

To test this prediction, we show on the right panel of Fig. 1, the exact Monte Carlo solution associated with the same barrier as before, including the void-in-cloud effect (red histogram), neglecting the void-in-cloud effect (black dotted), and the theoretical prediction of equation (9) (black solid line). As we can see, the agreement with the exact solution is quite accurate over the all range in radius. Note also that the correlations between steps induced by the SX filter decreases the number of voids by a non-negligible factor; thus, the effect is important and should be properly implemented in any void abundance prediction. Furthermore, the difference between the two-barrier model and the solution which neglects the void-in-cloud process is less important than for the SK case, a point first made by Paranjape et al. (2012). Similarly to the haloes cloud-in-cloud process, the physical reason why the void-in-void or void-in-cloud effects influence only small scales is that large-scale voids are most likely to be at the top of the hierarchy, not embedded in even larger voids or haloes. For SX filters, this Monte Carlo shows that the probability that an initial underdense patch of matter with Lagrangian radius $>1\,\mathrm{Mpc}\,h^{-1}$ is embedded in an overdense larger region is negligible. These results are also in agreement with Jennings et al. (2013). Note that the small influence of the void-in-cloud effect could also be due to the drifting terms which effectively increase the separation between the halo and void barriers. Overall those Monte Carlo tests show that equation (9) is a good prediction for the void abundance as long as the excursion set assumptions (e.g. averaging the smoothed field over random positions) can be applied to the description of void statistics and the void-in-cloud process is negligible. We describe an alternative peaks approach in the next section.

2.2 Peaks approach

In addition to their two-barrier SK random-walk model, SvdW also discussed alternative models based on counting density minima in the initial conditions. The model that they called 'adaptive troughs', which was based on previous work by Appel & Jones (1990) for the halo mass function, is especially interesting for us, because recent work on the nature of random walks with correlated steps sheds new light on its interpretation.

The adaptive troughs model states that the void multiplicity function can be written by using the Bardeen et al. (1986, hereafter BBKS) result for counting density peaks/troughs and including the effect of a variable smoothing filter:

$$f(\sigma) = \frac{\mathrm{e}^{-\delta_\mathrm{v}^2/2\sigma^2}}{\sqrt{2\pi}\gamma}\frac{V}{V_*}\int_0^\infty \mathrm{d}x\, x F(x)\, p_\mathrm{G}(x-\gamma|\delta_\mathrm{v}|/\sigma; 1-\gamma^2). \tag{10}$$

Here $V = 4\pi R^3/3$ is the Lagrangian volume of the void, $p_\mathrm{G}(y-\mu;\Sigma^2)$ is a Gaussian in the variable y with mean μ and variance Σ^2, and γ and V_* are ratios of spectral integrals that appear when counting density peaks/troughs:

$$\gamma \equiv \sigma_1^2/(\sigma_0\sigma_2); \quad V_* \equiv (6\pi)^{3/2}\sigma_1^3/\sigma_2^3, \tag{11}$$

where

$$\sigma_j^2 = \int \frac{\mathrm{d}^3k}{(2\pi)^3} P(k)\, k^{2j}\, \mathrm{e}^{-k^2R_\mathrm{G}^2}. \tag{12}$$

The Gaussian smoothing scale R_G depends approximately linearly on the Lagrangian radius R and is discussed below.

The integral in equation (10) is over the peak curvature $x = -\nabla^2\delta/\sigma_2$, and involves the weighting function $F(x)$ given by

$$F(x) = \frac{1}{2}\left(x^3-3x\right)\left\{\mathrm{erf}\left(x\sqrt{\frac{5}{2}}\right)+\mathrm{erf}\left(x\sqrt{\frac{5}{8}}\right)\right\} + \sqrt{\frac{2}{5\pi}}\left[\left(\frac{31x^2}{4}+\frac{8}{5}\right)\mathrm{e}^{-5x^2/8} + \left(\frac{x^2}{2}-\frac{8}{5}\right)\mathrm{e}^{-5x^2/2}\right], \tag{13}$$

which is the result of integrating over peak shapes [equations (A14–A19) in BBKS]. While there is no closed form expression for the multiplicity (10), the integral involved is straightforward to compute numerically, and we also note that BBKS provide a very accurate analytical approximation in their equations (4.4), (4.5), (6.13), and (6.14).

Recently, Paranjape & Sheth (2012) pointed out, based on results obtained by Musso & Sheth (2012), that the multiplicity in equation (10) is an excellent approximation to the first-crossing distribution of the constant barrier $B = \delta_\mathrm{v}$ by *peak-centred* random walks with correlated steps. Moreover, as argued by Paranjape et al. (2012), accounting for the complications introduced by the fact that voids are identified in Eulerian rather than Lagrangian space does not lead to significant effects when dealing with walks that have correlated steps. In particular, Paranjape et al. (2012) showed (see their fig. 3) that the appropriate first-crossing distribution for Eulerian voids (under the assumption of spherical evolution) is indistinguishable from that of a single constant barrier of height δ_v for all but the smallest voids. In other words, taken together, the results of Paranjape et al. (2012) and Paranjape & Sheth (2012) suggest that equation (10) should be a good model of void abundance, if one expects voids to have formed near initial density minima.

There is a technical issue related to the choice of Gaussian filtering with scale R_G in defining the spectral integrals in equation (12). Ideally one would use top-hat (SX) filtering to define these integrals. However, in this case the identification of peaks for the CDM power spectrum becomes ill-defined since, e.g. σ_2 is no longer well defined. Gaussian filtering avoids this problem, and all results in BBKS assume this. In order to make the calculation consistent with the standard assumption of defining δ using top-hat filtering, Paranjape, Sheth & Desjacques (2013) proposed the following: to identify peaks/troughs, one can use spatial derivatives of the Gaussian-filtered density contrast $\delta_\mathrm{G}(R_\mathrm{G})$ so that $\sigma_1^2 = \langle (\nabla\delta_\mathrm{G})^2 \rangle$ and $\sigma_2^2 = \langle (-\nabla^2\delta_\mathrm{G})^2 \rangle$ are well defined. The *heights* of these density extrema, on the other hand, can be defined using the top-hat filtered $\delta_\mathrm{TH}(R)$. The connection between the two smoothing scales R_G and R follows by demanding $\langle \delta_\mathrm{G}|\delta_\mathrm{TH} \rangle = \delta_\mathrm{TH}$. Since δ_G and δ_TH are both Gaussian distributed, this amounts to requiring $\langle \delta_\mathrm{G}\delta_\mathrm{TH} \rangle = \langle \delta_\mathrm{TH}^2 \rangle = \sigma^2$. This can be solved numerically and, in practice, gives $R_\mathrm{G} \approx 0.46R$ with a slow variation. To be fully consistent, one must also redefine γ as

$$\gamma \to \gamma_\mathrm{m} = \sigma_{1\mathrm{m}}^2/(\sigma\sigma_2) = \frac{1}{\sigma\sigma_2}\int \frac{\mathrm{d}^3k}{(2\pi)^3} P(k)\, k^2\, \mathrm{e}^{-k^2R_\mathrm{G}^2/2}\tilde{W}(kR), \tag{14}$$

with $\tilde{W}(kR)$ the Fourier transform of the top-hat filter, and where σ is defined in equation (3). (In practice, the top-hat filtered σ^2 and Gaussian filtered σ_0^2 differ by at most ∼5 per cent or so.)

These results can also be extended to the case when the barrier relevant for void formation is stochastic and/or scale dependent. Following Paranjape et al. (2013), for a barrier of the form

$$B_\mathrm{v} = \delta_\mathrm{v} - \beta_\mathrm{pk}\sqrt{S}, \tag{15}$$

with the slope β_pk a stochastic quantity with distribution $p(\beta_\mathrm{pk})$ in general, the void multiplicity becomes

$$f(\sigma) = \int \mathrm{d}\beta_\mathrm{pk}\, p(\beta_\mathrm{pk})\frac{\mathrm{e}^{-(\delta_\mathrm{v}-\beta_\mathrm{pk}\sigma)^2/2\sigma^2}}{\sqrt{2\pi}\gamma_\mathrm{m}}\frac{V}{V_*} \times \int_{\beta_\mathrm{pk}\gamma_\mathrm{m}}^\infty \mathrm{d}x\,(x-\beta_\mathrm{pk}\gamma_\mathrm{m})F(x) \times p_\mathrm{G}(x-\beta_\mathrm{pk}\gamma_\mathrm{m}-\gamma_\mathrm{m}|\delta_\mathrm{v}|/\sigma; 1-\gamma_\mathrm{m}^2). \tag{16}$$

The resulting multiplicity is shown in Fig. 1 for different parameters. The dotted orange curve shows the prediction from equation (10) for the constant barrier $B_\mathrm{v} = \delta_\mathrm{v} = -2.7$. The dashed orange curve shows the effect of introducing a negative drift with constant slope $\beta_\mathrm{pk} = 0.5$ [formally, equation (16) with $p(\beta_\mathrm{pk}) = \delta_\mathrm{D}(\beta_\mathrm{pk}-0.5)$], while the orange solid curve shows equation (16) setting $p(\beta_\mathrm{pk})$ to be log normal with mean 0.5 and variance 0.25. We return to those choices of parameters later.

First of all observe that the standard peak-based 'adaptive troughs' prediction (orange dotted curve) leads to a higher amplitude for the abundance of voids compared to the standard excursion set approach with a diffusive drifting barrier (see also SvdW). In addition, introducing scatter in the barrier height in the peaks prediction also tends to increase the number of voids, but it is a subdominant effect compared to the negative drift which decreases the amplitude. The solid orange line implements both scatter and drift while the orange dashed line neglects the scatter. Finally we note that, interestingly, both predictions for stochastic barriers with a negative drift are close to each other. For comparison, we also show the original SvdW prediction as the dot–dashed curve in which the relation between the density variance S and Lagrangian radius R was computed using the SX filter. While this is the usual manner in which the SvdW result is used, we emphasize that doing so is technically inconsistent, since the derivation in SvdW assumed SK filtering. Note also that the effect of the SX filter on the void abundance is to decrease the total number of voids.

Before moving to the next section, we should mention that the cumulative void volume fraction $\mathcal{F}(R > R_\mathrm{min})$ can be computed

from equations (16) and (9) only in the regime where the void-in-cloud process is negligible (i.e. $R_{\rm min} > 1\,{\rm Mpc}\,h^{-1}$).

3 COMPARISON WITH *N*-BODY SIMULATIONS

In order to test our theoretical predictions, we measure void abundances from *N*-body simulations using the ZOBOV void finder (Neyrinck 2008), described in the next subsection. First we consider the DM field of the Dark Energy Universe Simulation (DEUS) *N*-body simulations,[3] described in Alimi et al. (2010), Courtin et al. (2011), and Rasera et al. (2010), without any particle subsampling. We use two box sizes, of length 162 and 648 Mpc h^{-1}, both with 1024^3 particles, realized using the RAMSES code (Teyssier 2002) for a ΛCDM model calibrated to *Wilkinson Microwave Anisotropy Probe* (*WMAP*) 5-year parameters $(\Omega_{\rm m}, \sigma_8, n_{\rm s}, h, \Omega_{\rm b}) = (0.26, 0.79, 0.96, 0.7, 0.0456)$. Secondly, we study the effect of subsampling the particles and using biased tracers to identify voids. For this purpose we use randomly selected DM particles equal to the total number of haloes in the simulations. To study the effect of the bias, we compare the resulting void function to the one obtained using the halo catalogues as tracers. In both boxes, haloes are identified using the friends-of-friends algorithm with linking length $b = 0.2$.

3.1 ZOBOV

The ZOBOV (ZOnes Bordering On Voidness) void finder (Neyrinck, Gnedin & Hamilton 2005; Neyrinck 2008, hereafter N08) is designed to be parameter free. ZOBOV uses the adaptive, parameter-free Voronoi tessellation to estimate the density (e.g. Schaap & van de Weygaert 2000) at every particle. A void is grown around each local density minimum particle using a watershed transform (e.g. Platen, van de Weygaert & Jones 2007): in an analogy to rain falling and flowing across a terrain, a particle p gets associated with a density minimum $p_{\rm min}$ if a particle-to-particle path on the tessellation down the steepest density gradients from p ends at $p_{\rm min}$.

ZOBOV forms a parameter-free partition of all particles into so-called 'zones', each of which is a watershed region flowing down into a single density minimum. It then returns a void catalogue, consisting of voids (sets of zones joined together) and subvoids. But we introduce two parameters to prune the raw void catalogue to something physically corresponding to our theoretical model, in which there are no subvoids, i.e. 'voids in voids' are not double counted.

To obtain a disjoint set of voids with boundaries that are likely not spurious, we apply the 'specifying a significance level' strategy described in N08 to the raw void catalogue. In this strategy, a boundary between two adjacent zones is declared to be real if the 'density contrast ratio', i.e. the ratio between the ridge density (the lowest density along the ridge separating the voids) and the density minima, exceeds a threshold corresponding to a 2σ (95 per cent) probability that a void did not arise from Poisson noise (N08).

We also introduce a threshold to the minimum density in each void to eliminate local density minima in high-density regions (which will occur by chance in a sufficiently well-sampled high-density structure). As shown by N08, removing voids with density contrast under the 2σ threshold will typically also remove these high-density voids, but to be sure about this, we apply a threshold at the minimum density found in the void, called the 'core density' (void with minimum density).

[3] www.deus-consortium.org

In theory, a spherical void has $\Delta_{\rm v} = -0.8$. Using ZOBOV a void is composed of several zones which have different densities. The total mean density of all zones is what we expect to correspond best to the critical $\Delta_{\rm v} = -0.8$. However, several zones can have much higher density than the core zone. Thus, we use values of the core threshold ≤ -0.8.

Unfortunately it is difficult to know a priori what value to use for this threshold. It can be calibrated by measuring abundances of density contrast ratios in Poisson point samples, but this describes its statistical, not physical, significance. If this density contrast ratio is used to judge voidness, the void catalogue best corresponding to a physical set of voids would likely differ based on the mass resolution, or sampling level. In what follows, we use both $\Delta_{\rm v}^{\rm min} = -0.9$ and -0.8 which is one common choice used in the literature (e.g. Lavaux & Wandelt 2012; Pisani et al. 2014; Chuen Chan, Hamaus & Desjacques 2014).

One might wonder why we do not cut directly on the (volume-weighted) average density within the void. This average density is easily computed from quantities in the void catalogue, but in fact it can be quite noisy. This is because a watershed transform does not give boundaries that necessarily correspond to density contours; all that is required for a particle to belong to a void is that the particle is up a steepest density gradient from a density minimum, so in fact haloes might be included at the edge of a void. Still, a void's average density from ZOBOV gives additional information about it, and we will use it below.

We use a non-periodic box cut from a larger (periodic) volume, which approximates the situation one might consider observationally. To deal with boundaries, we follow the typical approach done with the ZOBOV algorithm (Granett et al. 2008; Sutter et al. 2012): we surround the box with a dense set of border particles. In the void-finding step, we exclude any particle that has a border particle detected as a neighbour. Unlike (Sutter et al. 2012), however, we do not remove voids from the catalogue that could be rotated to intersect the boundary.

3.2 The abundance of voids in the DM density field and for biased tracers

The number of voids identified with ZOBOV is sensitive to the density of the tracers inside the box. Naturally, if the number density of particles is high, then smaller voids will be detected. ZOBOV is designed to have low sensitivity to the sampling level for large voids, but the boundaries of voids change slightly when adding or subtracting particles randomly, so large voids are sometimes not exactly preserved. Also, increasing the mass resolution in a CDM simulation adds small-scale power, so decreasing mass resolution is not necessarily the same as randomly removing particles. As a result of these issues, the void abundance function can change slightly when particles are subsampled. In Pisani et al. (2014) and Chuen Chan et al. (2014), the authors subsampled particles of the DM field to match the density of the Slaon Digital Sky Survey (SDSS). In this work, we present two extreme cases: we consider the full DM density field and a sample of it for which the DM density equals the density of haloes in the simulation.

For the full DM density field, we analysed three subcubes of size $(40.5\,{\rm Mpc}\,h^{-1})^3$ from the $(162\,{\rm Mpc}\,h^{-1})^3$ box and two subcubes of size $(162\,{\rm Mpc}\,h^{-1})^3$ from the $(648\,{\rm Mpc}\,h^{-1})^3$ box. The total number of DM particles within the 40.5 and the $162\,h^{-1}$ Mpc subboxes is $N \sim 1.6 \times 10^7$, which is about the limit the QHULL algorithm (external module of ZOBOV) can treat.

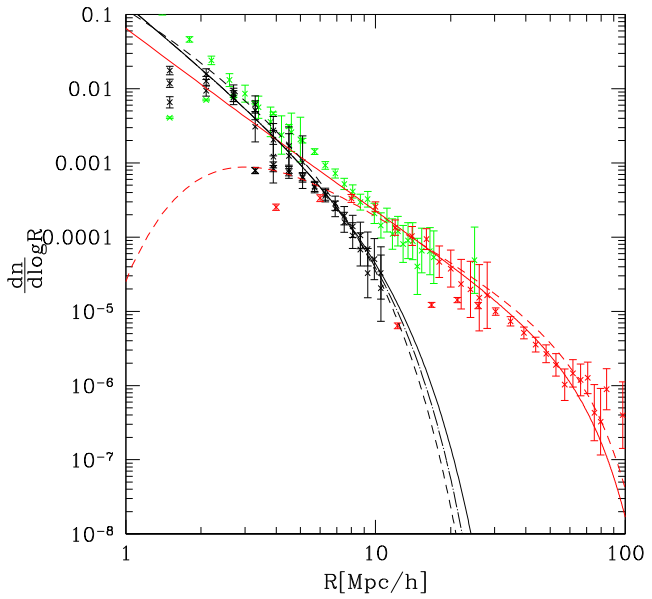

Figure 2. The void size distribution at $z = 0$ measured in simulations with different density tracers and choices of post-processing. The black and light green crosses are from the unsampled DM density fields and follow, respectively, case 1 and case 2 for the choice of post-processing (see text). The red crosses are from the haloes and follow case 2 for the choice of post-processing. Overplotted are the theoretical predictions of SvdW (dashed curves), DDB (solid curves), and peaks approach (dot–dashed curve).

To avoid counting the same structures twice, we remove subvoids by declaring boundaries to be real at a density contrast corresponding to a 2σ cut in a Poisson realization. This post-processing also removes spurious voids or subvoids arising from Poisson fluctuations in high-sampling limit. We used a 2σ cut as a compromise between having an accurate sample of disjoint voids (without subvoids) and a large number of voids. If we reduce the cut to 1σ, this would declare shallower boundaries separating subvoids within voids to be real, cutting up larger voids and shifting the distribution to small radius. Likewise, increasing the threshold removes walls between voids, which would increase the number of larger voids. We also try setting ZOBOV's core density threshold equal to -0.9. We define this choice of post-processing as case 1 (disregarding void boundaries under a 2σ significance cut, and using a -0.9 minimum density threshold). The corresponding results for void abundances are shown in Fig. 2 in black crosses. First three subcubes probe voids with radii $R \sim 1.5$–$5\,\mathrm{Mpc}\,h^{-1}$ while the other two subcubes probe voids with $R \sim 4.5$–$10\,\mathrm{Mpc}\,h^{-1}$.

Our second choice of post-processing is to put a minimum density threshold of -0.8 without a void boundary cut. We define this fiducial model as case 2. The result for case 2 on the unsampled DM density field is shown by the green crosses in Fig. 2 for two subcubes of size $(40.5\,\mathrm{Mpc}\,h^{-1})^3$ and $(162\,\mathrm{Mpc}\,h^{-1})^3$. In this case, the smaller subcube probes voids with radii $R \sim 1.5$–$5\,\mathrm{Mpc}\,h^{-1}$, while the larger subcube probes voids with $R \sim 4$–$18\,\mathrm{Mpc}\,h^{-1}$.

Note that in all cases, we removed from the voids list (output files of ZOBOV), those which have a large VoidDensContrast (>10). We further put a cut in the void function when the Poisson noise is higher than 50 per cent, for clarity of the figure. For both case 1 and case 2 we find a convergence of the void abundance within the different sub-boxes with different mass resolutions. Sensitivity to the choice of post-processing is significant for high-density tracers. This is because the voids from a high-density sample will contain a lot more subvoids and spurious voids compared to voids identified in low-density sample.

Unsurprisingly, using a 2σ cut of post-processing (case 1) reduces significantly the abundance of voids. The choice of core density threshold, -0.9, is rather arbitrary and is a compromise between matching the theory lines and choosing a threshold closer to what we can expect from a void with underdensity $\Delta = -0.8$ (a cut in the minimum density of a void should be lower than -0.8 to achieve a mean density in the void of -0.8).

To map the Lagrangian theory of equation (1) to effective void radius ($R = (3V/4\pi)^{1/3}$, where V is the void volume reported by ZOBOV), we use the spherical model and set $1.7R_{\mathrm{Lag}} = R$. One could adopt a different approach and use a different mapping which might arise from aspherical voids consistent with the Lagrangian underdensity (linear void threshold). We do not investigate this issue here. In the case of haloes, this issue is not very relevant. In fact, spherical overdensity halo finders are based on the non-linear spherical collapse model, although the linearly extrapolated spherical collapse threshold does not work in detail for predicting the halo mass function. Therefore we adopt the same pragmatic approach for voids. The black dashed curve in Fig. 2 shows the SvdW prediction with the usual $\delta_{\mathrm{v}} = -2.7$ and $\delta_{\mathrm{c}} = 1.68$. The black solid curve uses equation (9), while the black dot–dashed curve uses equation (16), with the same parameters as in Fig. 1.

All the black theory curves match the data reasonably well. The parameters for the barrier we adopt were, strictly speaking, motivated by parameters for the halo, not void, mass function (Achitouv et al. 2013; Paranjape et al. 2013). We do not expect the same criteria to hold for both, although this is a good starting point. The agreement is a nice result, since equations (9) and (16) correspond to the prediction of the SX filter, which assumes a spherical volume to map the variance to the radius of the voids; this is one of the main differences compared to the SvdW prediction. Note that, in contrast to the voids identified by Jennings et al. (2013), the SvdW curve is quite close to the volume function of our ZOBOV voids, with no factor of $\sim$5 offset. In fact, redefining voids as proposed by Jennings et al. (2013), might be more consistent with the spherical shell approximation. In order to defined voids as 'spherical underdensities', Jennings et al. (2013) use density minima found by ZOBOV to get initial void and subvoid centres, but then they define the boundaries quite differently: for each void and subvoid, they start from a large radius about the centre, and decrease the radius until the underdensity reaches $\Delta_{\mathrm{v}} = -0.8$. While this procedure intuitively corresponds well to a spherical shell approximation, the voids can be quite aspherical, so this procedure is not guaranteed to give the optimal correspondence with theory. In full non-linearity, voids run into each other and their boundaries typically depart substantially from a spherical shape; these arbitrary shapes will be picked up by ZOBOV. Note also that Jennings et al. (2013) compare the 'spherical underdensity' voids with the excursion set theory associated with the SK filter.

Finally, we investigate the effect of sampling the tracers to identified voids. In fact, for very low density tracers, we expect void catalogues to contain large voids, while for very dense tracers (as for the full DM density field at the resolution here), voids are rather small. Therefore we probe large voids by running ZOBOV on the halo positions of our two simulations with a core density threshold -0.8 and no σ cut (case 2). The halo densities in our two boxes are $\sim$0.073 and $\sim 0.001\,h^3\,\mathrm{Mpc}^{-3}$ for, respectively, the 162 and 648 $\mathrm{Mpc}\,h^{-1}$ box sizes. The result is shown in Fig. 2 by the red crosses. The $162\,\mathrm{Mpc}\,h^{-1}$ simulation probes void between $\sim$10 and $20\,\mathrm{Mpc}\,h^{-1}$ while the $648\,\mathrm{Mpc}\,h^{-1}$ simulation probes voids

between ∼20 and 80 Mpc h^{-1}. Again, we observe a good overlap of the void abundance between the different simulations.

In this low-density sample, our choice of post-processing would not change significantly the abundance of voids. In Chuen Chan et al. (2014) the authors observed that if the density of the tracers is not high enough, smaller voids tend to artificially merge to form larger voids. To avoid this effect, a statistical threshold like the one we chose for the unsampled DM particle is not enough. Nevertheless, as we can see in Fig. 2, the abundance of these voids can be fitted by adjusting the value of δ_c which enters into the theoretical prediction. The red dashed curve shows the theory of SvdW using $\delta_v = -0.6$. The red solid curve shows the DDB equation (9) result for $D_B = 0.4, \beta = -0.1$, and $\delta_v = -0.6$. As we can see, once again we obtain a good agreement with the data. One could ask if the parameters $\delta_v = -0.6$ can be related to any measurement within the simulation. This goes beyond the scope of this paper since it would require an investigation of the mapping between the size of the void and the protovoid. This mapping is well established in the case of spherical evolution ($1.7R_{Lag} \sim R$). Hence for the void catalogue fitted by our theoretical formula with the spherical barrier prediction $\delta_v = -2.7$, we will test further the consistency of the spherical evolution in Section 4. Before we will briefly investigate the effect of the biasing over the identification of voids.

Finally, unlike the DDB model which is quite flexible on its own, we have found that the ESP model cannot describe the size distribution of voids identified using haloes (or sparse DM samples). This is perhaps not so surprising; discrete tracers of voids would likely need to be addressed using a full halo model analysis (e.g. Furlanetto & Piran 2006). This is beyond the scope of the paper and we leave such an investigation to future work.

3.3 The effect of biasing

Previously we considered halo positions to study the effect of subsampling matter particles on the abundance of voids. Possibly, biasing could change the void volume function from that of the DM. To see how, we run ZOBOV under the same void selection (core density −0.8) that we previously chose for the halo tracers. We randomly selected DM particles with a total number equal to the number of haloes, for each box. In Fig. 3 we can see the resulting void abundance function: red crosses correspond to haloes while blue crosses correspond to DM particles for the tracers. As we can see, the effect of biasing is negligible on the abundance of small voids. The abundance of large voids reduces slightly when they are identified by the DM particles. For comparison, the blue dashed curve shows the SvdW prediction for $\delta_v = -0.7$, the blue solid curve is the DDB model with ($D_B = 0.4, \delta_v = -0.7, \beta = -0.2$).

To conclude this section, we have shown the difference in the void abundance when the tracer density changes. This allowed us to probe a large range of void radii. We find that without additional post-processing, the $dn/d\log R$ function can be fitted by the SvdW and our theory equation (9) once we rescale $\delta_v = -2.7$ to ~ -0.6. We note that without further post-processing, a fraction of these voids have an underdensity above zero and are strictly speaking not real voids. We find that biased tracers have an effect on large voids. Finally, we recover an agreement between the data and theoretical predictions of the void abundance with the usual spherical criteria $\delta_v = -2.7$ for the choice of post-processing case 1, described in the previous section. For the voids thus identified, we next investigate the linear critical underdensities that lead to their formation.

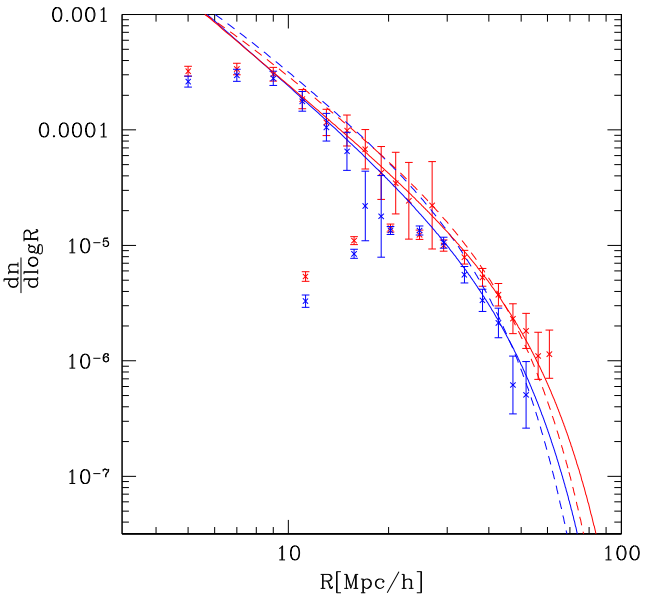

Figure 3. Effect of biasing on the void abundance: blue crosses correspond to void identified in the DM density field and red crosses correspond to void identified using haloes as biased tracers. The theoretical prediction of SvdW (dashed curves) and DDB (solid curves) is shown for different values of the barrier parameters (see text).

4 CONSISTENCY OF THE SPHERICAL EVOLUTION APPROXIMATION

A straightforward way to relate non-spherical voids to the spherical collapse model is to define them around density minima as spheres with underdensity $\Delta_v = -0.8$, as in Jennings et al. (2013). Then for all of those voids, we could go back in the initial conditions and test whether the protovoid corresponds to a linear underdensity of $\delta_v = -2.7$. However, as we already mentioned, including the edge of the void or not makes a significant difference for the void's average density. Therefore, defining voids through a density criterion might be noisy. In this sense, ZOBOV may be more suitable to define voids and compare with observation. Nevertheless, from the theoretical modelling, the spherical shell evolution is generally assumed, and in this section we propose to test whether this assumption is consistent or not. In order to perform this test, we first consider the $z = 0$ void volume centroid, a Voronoi volume-weighted average of particle positions. This centre can differ from the one defined using the particle which sits on the minimum density of the void (the 'core particle'). Indeed, the minimum of the density profile might not correspond to the minimum of the potential if the surrounding shear field is asymmetric. Therefore for each void, we find the closest particle to its volume centroid. Going back to the initial conditions, we record the underdensity in a sphere centred on this centre particle within the Lagrangian radius of the corresponding void.

In the ideal spherical shell model, the distribution of this critical underdensity[4] $\delta \equiv \delta_x$ for all void size $\Pi(\delta_x, S(R))$ would be a Dirac delta centred on $\delta_v = -2.7$. Indeed, spherical evolution is fully deterministic. Following the procedure we just described, the probability density function (PDF) we measure is shown in Fig. 4 for two different void sizes.

[4] The x subscript means that the barrier crosses the smoothed overdensity: $\delta_x \equiv \{B_v \cap \delta\}$.

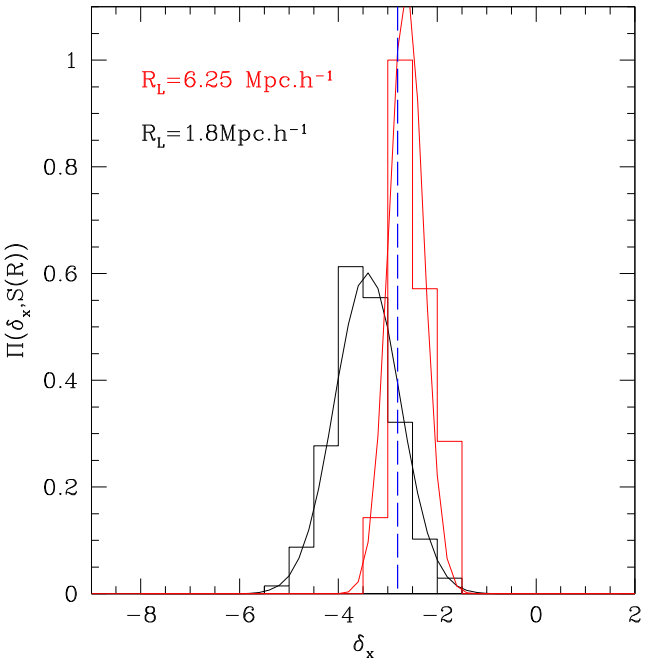

Figure 4. Initial overdensities (linearly extrapolated to $z = 0$) measured in the N-body simulations of regions which produce voids at $z = 0$ for two different void radii. Solid lines correspond to Gaussian fits. Blue vertical line shows the spherical model prediction.

First of all, note that the initial overdensities in Fig. 4 are negative as expected, which confirms that voids form from underdense regions. However, there is a significant scatter, indicating that the spherical threshold cannot be exactly applied to model the abundance of voids defined as in our previous analysis. Secondly, the mean of the distribution varies for two different voids sizes ($\langle\delta_x\rangle(R = 1.8) = -3.4$, $\langle\delta_x\rangle(R = 6.25) = -2.5$). While the mean value is close to the spherical prediction for the larger void size, at smaller sizes the implied threshold is significantly more underdense, with a larger scatter. In principle, the scale dependence of the mean value can be modelled by a negative drift term. Note that this analysis only assumes spherical dynamics on an object-by-object basis; we do not assume any particular statistical model for the void abundance.

In the peaks-based model this would imply a mean value $\langle\,\beta_{\rm pk}\,\rangle \simeq 1$ when using equation (15). More precisely, given the prior $p(\beta_{\rm pk})$ and the peaks calculation of $f(\sigma|\beta_{\rm pk})$, we calculate the volume function $f(\sigma) = \int {\rm d}\beta_{\rm pk} p(\beta_{\rm pk}) f(\sigma|\beta_{\rm pk})$ as described in Section 2.2. Then we use Bayes' rule to calculate $p(\beta_{\rm pk}|\sigma) = f(\sigma|\beta_{\rm pk})p(\beta_{\rm pk})/f(\sigma)$. This is directly related to $p(\delta_x|\sigma)$ upon using equation (15).

For the DDB model, the value of $\beta, D_{\rm B}$ can be inferred using the mapping of equation (10) in Achitouv et al. (2013) once we build the PDF of δ_x at randomly selected positions. Note, however, that different sets of values would not lead to a good agreement with the measured volume function, indicating a possible breakdown of the spherical evolution assumption which we discuss further below.

In addition, we checked that the closest particle to the centre void at $z = 0$ is at a distance d well below the void radius R, to avoid a bias due to a wrong definition of the void centre. Also, we checked that the PDF of δ_x is almost insensitive to the displacement of this centre particle. This is important because if we assume that this particle sits at the minimum of the density, then tracking its displacement tells us if the void forms from an initial underdense peak in Lagrangian space. It also reassures us that we have picked the correct centre of the void: if initially this particle is not at the minimum of the potential then it should be pulled out of the centre at $z = 0$.

Note that alternatively to our protovoid centre definition, we could have defined the centre in Lagrangian space by averaging over initial positions of particles which belong to the void at $z = 0$, weighting each particle in the average by its $z = 0$ Voronoi volume. Since most of the uncertainty comes from the void edge, where densities are high; it may be preferable to downweight particles there. However, we expect this definition of the void centre to lead to nearly the same results in Fig. 4. Indeed, we found that the PDF in Fig. 4 changes only slightly when the void centres used to measure densities in the initial conditions are displaced to their final positions. This suggests that Fig. 4 should be insensitive to subtleties in void centring.

To summarize, this analysis shows that, as expected, voids form from initial underdense regions (see Fig. 4). Most of the void centres displace over time from their initial location, but the density criterion which leads to their formation is rather uncorrelated to this displacement. Furthermore, the deterministic spherical evolution is apparently not exactly achieved even for larger voids. This could indicate a breakdown of the simple spherical model. It is possible, however, that a modification to the void finder would improve agreement with the model; to test for that, we perform an additional analysis in the next section, selecting spherical underdensity voids with top-hat average densities $\Delta_{\rm v} = -0.8$ at $z = 0$.

5 VOIDS DEFINED AS SPHERICAL UNDERDENSITIES

Previously, we tested if ZOBOV voids have initial underdensities in agreement with the spherical threshold. Furthermore, the spherical evolution of an isolated underdense patch, leading to a void at $z = 0$, gives specific predictions (SvdW). Assuming

a – initial (proto) voids are spherical top-hats;

this spherical protovoid evolves similarly to an Friedmann–Lemaître–Robertson–Walker (FLRW) universe. At shell crossing,

b – initial voids enclose a linear density contrast corresponding to $\delta_{\rm v} = -2.71$ at $z = 0$;

c – voids at $z = 0$ have a density contrast $\Delta_{\rm v} = -0.8$; and

d – voids at $z = 0$ are spherical.

This relation between the linear and non-linear underdensity of the spherical shell predicts

e – the radius at $z = 0$ is larger by a factor $q = 1.7$ compared to the initial radius ($R_{\rm E} \sim qR_{\rm L}$).

In this section, we select voids in our catalogues which satisfy points (d) and (c), and test the rest of these predictions. For this purpose, we follow the approach of Jennings et al. (2013). We start from ZOBOV void centres (density minima in the Voronoi tessellation) at $z = 0$. Then we measure the density at large radius, and move void radius inward until $\Delta_{\rm v} = -0.8$. We find that the radius tends to be smaller than the one we find from the ZOBOV output which does not impose a density threshold. Typically the difference is of order $R_{\rm E} = 0.9R_{\rm E}^{\rm ZOBOV}$.

Then we repeat the previous analysis: assuming points (a) and (e) are satisfied, we compute the linear underdensities leading to those voids. In this case, we should recover point (b). Fig. 5 shows the result of this new analysis, at two different void sizes. As we can see, the linear underdensity is again not deterministic but it has a

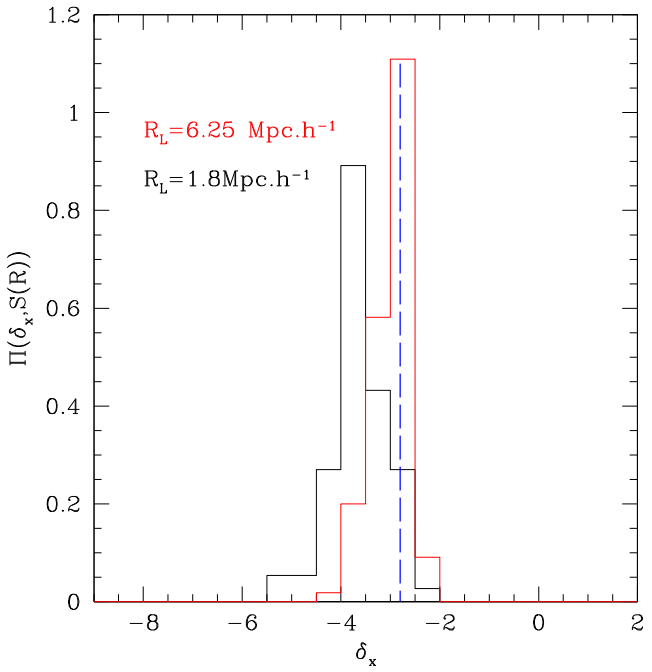

Figure 5. Initial overdensities (linearly extrapolated to $z = 0$) measured in the *N*-body simulations of regions which produce spherical voids of $\Delta_v = -0.8$ at $z = 0$ for two different void radii. Blue vertical line shows the spherical model prediction.

distribution which varies as a function of the void size. We report a mean value $\langle \delta_x(R = 1.8) \rangle = -3.6$ and $\langle \delta_x(R = 6.25) \rangle = -2.9$. These are smaller than the previous means (-3.4 and -2.5, respectively). This decrease is not surprising since on average, the new voids are smaller, comprising mostly the cores of ZOBOV voids. For large voids, the variance is reduced compared to that in Fig. 4, indicating that spherical underdensity voids follow a deterministic evolution better than ZOBOV voids. There is a decreased variance compared to that in Fig. 4 for the smaller voids as well, although the mean value is still smaller than the spherical model expectation especially for small voids. In the model of Paranjape et al. (2012), the authors find that the Lagrangian patch corresponding to a void would have a (linearly extrapolated) density contrast below -2.7 as long as its surrounding has a density to restrict the expansion (hence forming a wall outside of the void). Since small voids are more likely to be affected in that model, the average density contrast smoothed over the Lagrangian scale would be smaller than those of the large voids, matching the trend observed in the simulation.[5]

Note that one could probably change point (e) such that q is a function void size, giving $\delta_v = -2.7$ for all voids. Typically, for small voids, this factor would have to be bigger in order to enclose more matter in the initial conditions. However, this would not remove the associated scatter around the mean value. There are a couple of factors that could contribute to this scatter. First, the particle nearest the volume centroid does not necessarily occupy a density minimum in the initial conditions even for an isolated void; it may even be at a small halo. However, such a peak corresponding to a small halo is unlikely to persist in the initial conditions after top-hat smoothing with the Lagrangian void radius. Indeed, we did test that using an alternative definition of a void centre (the density minimum, instead of volume centroid) produced nearly identical results. Secondly, interactions between voids could lead to a poor correspondence between $z = 0$ voids and initial density minima. For example, if two voids that accurately follow spherical evolution are very close to each other, they could merge together, with an undetectably tenuous density ridge between them. Possibly, both ZOBOV and the spherical underdensity algorithm could report a point near the ridge between them as the void centre. This would give a flawed estimate of the radius of the void; also, the reported centre could fail to be an underdensity in the initial conditions.[6]

Despite these possible issues, Jennings et al. (2013) show that this procedure leads to reasonable predictions of void abundances. Overall it is quite remarkable that the simple spherical model shows an approximate consistency for the large voids, for both ZOBOV and spherical underdensity voids.

6 CONCLUSIONS

In this paper we develop simple analytical predictions for the void volume function based on the excursion set approach and the peaks formalism which consistently include the effects of filtering with a top-hat in real space. We extended traditional predictions that use a deterministic barrier threshold to a more general class of barriers which are stochastic and sensitive to the size of the voids. The analytical predictions were tested against Monte Carlo simulations of random walks crossing the diffusive drifting barrier and show very good agreement.

We also compared our analytical predictions for the void volume function with numerical results obtained by running the ZOBOV void finder on *N*-body simulations with different mass resolution and choice of post-processing.

ZOBOV is particularly suited to find highly aspherical voids in the cosmic web, as many underdensities are at the present epoch. One (sometimes underappreciated) aspect of ZOBOV is that the measured void volume function can depend on the choices made on the core threshold parameter. Furthermore, many low-density contrast voids arise quite frequently in a Poisson process and are likely to be spurious, not corresponding to a physical void. Removing these voids by performing an additional post-processing can change the trend of the void volume function. While this should be kept in mind when analysing observational results based on ZOBOV this sensitivity comes in only in the high sampling limit (when discreteness noise produces potentially spurious voids).

Regardless of post-processing analysis, we find a reasonable agreement between the analytical predictions (SvdW; DDB) and the measured void volume functions once we arbitrary changed the barrier parameters. For one specific choice of post-processing we found that the SvdW, DDB, and the peak models reproduced the measured void volume function with the traditional value of $\delta_v = -2.7$. However, the specific parameter values of the stochastic barrier needed to obtain this agreement (at least for the peaks-based model) do not appear to be consistent with our direct measurements of the overdensity threshold in the initial conditions. In fact, we showed in a direct measurement in the initial conditions that the usual linear theory threshold is incompatible with the deterministic value predicted by the spherical model for these ZOBOV voids. These voids have an effective radius less than $10\,h^{-1}$ Mpc at $z = 0$. This is an important result since current predictions for the abundances of voids defined using compact objects require the linear threshold criterion as an input (Furlanetto & Piran 2006). We also tried

[5] We are grateful to the referee for pointing out this explanation.

[6] We thank the referee for pointing out this second effect.

a different, spherical underdensity void finder, as Jennings et al. (2013) used. With this approach too we find that small voids do not obey spherical evolution in detail; their linearly extrapolated initial densities are generally lower than spherical evolution would predict while larger voids tend to agree with the spherical threshold in average.

Our work can be extended in several directions: one could compare our prediction with a spherical underdensity algorithm and see if the corresponding initial threshold tends to the deterministic spherical prediction, introduce a more general barrier as we present here, or perhaps an deeper investigation on the mapping between Lagrangian and Eulerian space could be performed by empirically finding a q factor ($R/q = R_{\rm Lag}$) such that the protovoid underdensity distribution becomes consistent with the linear barrier which enters in the modelling of the void abundance. If such a factor can be found we can already infer that it would have to be smaller than the spherical 1.7 value in order to get a shallower mean underdensity for small voids. Those tests are crucial if we want to have analytical model predictions able to compare to current and future observational void catalogues.

ACKNOWLEDGEMENTS

We thank Yann Rasera for useful discussions about the DEUS *N*-body simulation, and Ben Wandelt for his reception at IAP. IA acknowledges support from the Trans-Regional Collaborative Research Center TRR 33 'The Dark Universe' of the Deutsche Forschungsgemeinschaft (DFG). MN is grateful for support from a New Frontiers of Astronomy and Cosmology grant from the Sir John Templeton Foundation.

REFERENCES

Achitouv I. E., Corasaniti P. S., 2012a, J. Cosmol. Astropart. Phys., 2, 2
Achitouv I. E., Corasaniti P. S., 2012b, Phys. Rev. D, 86, 083011
Achitouv I., Rasera Y., Sheth R. K., Corasaniti P. S., 2013, Phys. Rev. Lett., 111, 231303
Achitouv I., Wagner C., Weller J., Rasera Y., 2014, J. Cosmol. Astropart. Phys., 10, 77
Alimi J.-M., Füzfa A., Boucher V., Rasera Y., Courtin J., Corasaniti P.-S., 2010, MNRAS, 401, 775
Appel L., Jones B. J. T., 1990, MNRAS, 245, 522
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15 (BBKS)
Bertschinger E., 1985, ApJS, 58, 1
Biswas R., Alizadeh E., Wandelt B. D., 2010, Phys. Rev. D, 82, 023002
Blumenthal G. R., da Costa L. N., Goldwirth D. S., Lecar M., Piran T., 1992, ApJ, 388, 234
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Chuen Chan K., Hamaus N., Desjacques V., 2014, Phys. Rev. D, 90, 103521
Colberg J. M., Sheth R. K., Diaferio A., Gao L., Yoshida N., 2005, MNRAS, 360, 216
Colberg J. M. et al., 2008, MNRAS, 387, 933
Corasaniti P. S., Achitouv I., 2011a, Phys. Rev. D, 84, 023009
Corasaniti P. S., Achitouv I., 2011b, Phys. Rev. Lett., 106, 241302
Courtin J., Rasera Y., Alimi J.-M., Corasaniti P.-S., Boucher V., Füzfa A., 2011, MNRAS, 410, 1911
Croton D. J. et al., 2004, MNRAS, 352, 828
D'Aloisio A., Furlanetto S. R., 2007, MNRAS, 382, 860
D'Amico G., Musso M., Noreña J., Paranjape A., 2011, Phys. Rev. D, 83, 023521
Dubinski J., da Costa L. N., Goldwirth D. S., Lecar M., Piran T., 1993, ApJ, 410, 458
Epstein R. I., 1983, MNRAS, 205, 207
Fillmore J. A., Goldreich P., 1984, ApJ, 281, 9
Furlanetto S. R., Piran T., 2006, MNRAS, 366, 467
Goldberg D. M., Jones T. D., Hoyle F., Rojas R. R., Vogeley M. S., Blanton M. R., 2005, ApJ, 621, 643
Granett B. R., Neyrinck M. C., Szapudi I., 2008, ApJ, 683, L99
Hamaus N., Wandelt B. D., Sutter P. M., Lavaux G., Warren M. S., 2014, Phys. Rev. Lett., 112, 041304
Hoffman G. L., Salpeter E. E., Wasserman I., 1983, ApJ, 268, 527
Hoyle F., Vogeley M. S., 2002, ApJ, 566, 641
Hoyle F., Vogeley M. S., 2004, ApJ, 607, 751
Hoyle F., Rojas R. R., Vogeley M. S., Brinkmann J., 2005, ApJ, 620, 618
Hunt P., Sarkar S., 2010, MNRAS, 401, 547
Jennings E., Li Y., Hu W., 2013, MNRAS, 434, 2167
Kamionkowski M., Verde L., Jimenez R., 2009, J. Cosmol. Astropart. Phys., 1, 10
Kauffmann G., Fairall A. P., 1991, MNRAS, 248, 313
Kirshner R. P., Oemler A., Jr, Schechter P. L., Shectman S. A., 1981, ApJ, 248, L57
Lacey C., Cole S., 1993, MNRAS, 262, 627
Lam T. Y., Sheth R. K., 2009, MNRAS, 398, 2143
Lam T. Y., Sheth R. K., Desjacques V., 2009, MNRAS, 399, 1482
Lavaux G., Wandelt B. D., 2010, MNRAS, 403, 1392
Lavaux G., Wandelt B. D., 2012, ApJ, 754, 109
Maggiore M., Riotto A., 2010a, ApJ, 711, 907
Maggiore M., Riotto A., 2010b, ApJ, 717, 515
Melchior P., Sutter P. M., Sheldon E. S., Krause E., Wandelt B. D., 2014, MNRAS, 440, 2922
Musso M., Sheth R. K., 2012, MNRAS, 423, L102
Neyrinck M. C., 2008, MNRAS, 386, 2101 (N08)
Neyrinck M. C., Gnedin N. Y., Hamilton A. J. S., 2005, MNRAS, 356, 1222
Pan D. C., Vogeley M. S., Hoyle F., Choi Y.-Y., Park C., 2012, MNRAS, 421, 926
Paranjape A., Sheth R. K., 2012, MNRAS, 426, 2789
Paranjape A., Lam T. Y., Sheth R. K., 2012, MNRAS, 420, 1648
Paranjape A., Sheth R. K., Desjacques V., 2013, MNRAS, 431, 1503
Park D., Lee J., 2007, Phys. Rev. Lett., 98, 081301
Patiri S. G., Betancort-Rijo J., Prada F., 2006a, MNRAS, 368, 1132
Patiri S. G., Betancort-Rijo J. E., Prada F., Klypin A., Gottlöber S., 2006b, MNRAS, 369, 335
Peacock J. A., Smith R. E., 2000, MNRAS, 318, 1144
Pisani A., Lavaux G., Sutter P. M., Wandelt B. D., 2014, MNRAS, 443, 3238
Platen E., van de Weygaert R., Jones B. J. T., 2007, MNRAS, 380, 551
Press W. H., Schechter P., 1974, ApJ, 187, 425
Rasera Y., Alimi J.-M., Courtin J., Roy F., Corasaniti P.-S., Füzfa A., Boucher V., 2010, in Alimi J.-M., Fuözfa A., eds, AIP Conf. Ser. Vol. 1241, Invisible Universe. Am. Inst. Phys., New York, p. 1134
Robertson B. E., Kravtsov A. V., Tinker J., Zentner A. R., 2009, ApJ, 696, 636
Ryden B. S., 1995, ApJ, 452, 25
Sahni V., Sathyaprakah B. S., Shandarin S. F., 1994, ApJ, 431, 20
Schaap W. E., van de Weygaert R., 2000, A&A, 363, L29
Seljak U., 2000, MNRAS, 318, 203
Sheth R. K., van de Weygaert R., 2004, MNRAS, 350, 517 (SvdW)
Sutter P. M., Lavaux G., Wandelt B. D., Weinberg D. H., 2012, ApJ, 761, 44
Teyssier R., 2002, A&A, 385, 337
Tinker J. L., Conroy C., 2009, ApJ, 691, 633
van de Weygaert R., Platen E., 2011, Int. J. Modern Phys. Conf. Ser., 1, 41
van de Weygaert R., van Kampen E., 1993, MNRAS, 263, 481
Zhang J., Hui L., 2006, ApJ, 641, 641

This paper has been typeset from a TEX/LATEX file prepared by the author.